

Setting Rules for 2.7 Billion: a (First) Look into Facebook's Norm-Making System; Results of a Pilot Study

Kettemann, Matthias C.; Schulz, Wolfgang

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:

Kettemann, M. C., & Schulz, W. (2020). *Setting Rules for 2.7 Billion: a (First) Look into Facebook's Norm-Making System; Results of a Pilot Study*. (Working Papers of the Hans-Bredow-Institut | Works in Progress, 1). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI). <https://doi.org/10.21241/ssoar.71724>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/3.0>



LEIBNIZ INSTITUTE
FOR MEDIA RESEARCH
HANS-BREDOW-INSTITUT

Matthias C. Kettemann / Wolfgang Schulz

SETTING RULES FOR 2.7 BILLION

A (First) Look into Facebook's Norm-
Making System: Results of a Pilot Study

Hamburg, January 2020



Kettemann, Matthias C.; Schulz, Wolfgang: Setting Rules for 2.7 Billion. A (First) Look into Facebook's Norm-Making System: Results of a Pilot Study. Hamburg: Working Papers of the Hans-Bredow-Institut | Works in Progress # 1, January 2020.

ISSN 1435-9413 , DOI: <https://doi.org/10.21241/ssoar.71724>

CC BY NC ND 3.0



The series "Working Papers of the Hans-Bredow-Institut" [Arbeitspapiere des Hans-Bredow-Instituts] is published on an irregular basis at Verlag Hans-Bredow-Institut. The single issues can be downloaded as pdf-files from the Institute's website: https://hans-bredow-institut.de/en/publications/working_papers A printout of the issue is available directly from the publisher for a fee of 20.00 EUR.

Leibniz Institute for Media Research | Hans Bredow Institute (HBI), Hamburg

The research perspective of the Leibniz-Institute for Media Research | Hans-Bredow-Institut focuses on media transformation and related structural changes of public communication. With its cross-medial, interdisciplinary and independent research, it combines basic research and transfer research, and thus, generates knowledge on issues relevant for politics, commerce and civil society. This kind of media research requires cooperation: international comparative questions are dealt with with partners in many countries. More at www.leibniz-hbi.de.

The authors

Prof. Dr. Wolfgang Schulz is Director of the Leibniz Institute for Media Research | Hans Bredow Institute in Hamburg, Professor for "Media Law and Public Law including its Theoretical Foundations" at the Faculty of Law at the University of Hamburg and also Director of the Alexander von Humboldt Institute for Internet and Society in Berlin

PD Dr. Matthias C. Kettemann, LL.M. (Harvard) is head of the research programme "Regulatory Structures and the Emergence of Rules in Online Spaces" at the Leibniz Institute for Media Research and Visiting Professor for Questions on the Foundation and Future of the Rule of Law at the University of Heidelberg.

Leibniz Institute for Media Research | Hans Bredow Institute (HBI)

Rothenbaumchaussee 36

20148 Hamburg

Germany

Phone: (+49 40) 45 02 17-0

info@leibniz-hbi.de

www.leibniz-hbi.de





TABLE OF CONTENTS

Abstract	5
1. Introduction.....	7
2. Methodological Approach.....	12
3. Conceptualizing Governance: How Facebook Approaches the Policy Development Process	15
3.1. Constellations of Actors.....	15
3.2. Shared Frames of Relevance.....	16
3.3. Patterns of Communicative Practices.....	21
4. Doing Governance: Case Study-based Analyses	23
4.1. Constellation of Actors	23
4.2. Shared Framing as Normative Work.....	24
4.3. Communicative Practices	26
5. Discussion: Rule Production as Ordering.....	28
5.1. "Making Rules Up"	28
5.2. Constellations of Actors: Aiming for Throughput Legitimacy.....	28
5.3. Frame of Relevance: Autonomously Constructing a Legal Order	30
5.4. Communicative Practices: Planned Engagement	32
5.5. Outlook.....	33



ABSTRACT

This paper presents the outcome of a pilot study into the private order of communication developed by a major social network provider, Facebook Inc., and its policy development process. It is part of a broader research focus on the evolution and application, the legitimacy and contestation of norms in private online communication spaces and their public impact, both in terms of individual rights and societal cohesion. With first-of-a-kind access to the internal processes of policy development (norm production), researchers were able to study the development of content-related policies through phases of participant observation, expert interviews, and normative analyses. The first insights developed in this case study already make an important contribution to the understanding of the challenges posed by creating private rules for what are essentially global digital communication spheres, the interactions between rule-making processes within and outside Facebook, Inc., as a popular social media company that sets rules for 2.7 billion users, and the (self)-conception (and production) of legitimacy in norm-development through proceduralization and external stakeholder involvement. Facebook, we find in this case study, is developing its own normative order; its norms (community standards) are closely intertwined with its platform. It is the Product Policy team that is involved in developing norms. This is no accident. National legal systems need to be more intricately connected to the diversified (and still diversifying) order(s) of private communication.

Keywords: private ordering, terms of service, community standards, public communication, policy development, human rights, Facebook, legitimacy, freedom of expression, normative orders.

1. INTRODUCTION*

Humans are socio-communicative beings. We love to communicate. We construct our identities through our communicative relationships. With the widespread use of the Internet, the spaces in which these processes take place have, *first*, progressively been *digitalized*. As the European Court of Human Rights noted in 2015, “the Internet has now become one of the principal means by which individuals exercise their right to freedom to receive and impart information and ideas, providing [...] essential tools for participation in activities and discussions concerning political issues and issues of general interest.”¹ It plays “a particularly important role with respect to the right to freedom of expression.”²

Second, socio-communicative spaces have not only been enriched by taking on a digital dimension, they have also changed in character: a majority of online communication takes place in privately owned and regulated communicative settings. The key questions regarding how to enable, moderate and regulate speech today have to be asked and answered with a view to *digital* and *private* spaces. These changes in communicative spatiality take nothing away from the primary responsibility and ultimate obligation of states to protect human rights and fundamental freedoms, online just as offline.³ However, over the last decade, tension between the normativity inherent in the role of states and the facticity of online communicative practices that are both *de facto* and *de jure* as well as being primarily regulated by the rules of private actors is increasing in intensity. This becomes evident in court cases,⁴ which show marked jurisdictional divergences.

* Research for this paper has been conducted within the framework of Research Program 2: Regulatory Structures and the Emergence of Rules in Online Spaces (headed by Matthias C. Kettemann and Jan-Hinrik Schmidt) and the leading project Doing Internet Governance: Constructing Normative Structures Inside and Outside Intermediary Organizations (Lead: Wolfgang Schulz and Matthias C. Kettemann) at the Leibniz Institute for Media Research | Hans-Bredow-Institut (HBI), Hamburg. Principal investigators: Wolfgang Schulz and Matthias C. Kettemann. The methodology of the project has benefited from discussions with Uwe Hasebrink, Claudia Lampert, Wiebke Loosen, Jan-Hinrik Schmidt, Stephan Dreyer, Sascha Hölig, Malte Ziewitz and Irene Broer. Matthias C. Kettemann conducted the underlying empirical research. This paper was written equally by Matthias C. Kettemann and Wolfgang Schulz. Anna Sophia Tiedeke provided valuable research support. Jan-Hinrik Schmidt and Stephan Dreyer provided substantive comments. The authors express their gratitude to Facebook, especially Peter Stern, for giving our researcher unprecedented access to the Product Policy team and enabling participation in team and individual meetings both at Facebook HQ and over the Internet. No other support such as funding was requested or provided to the research team by Facebook. The project was completely self-funded by the HBI. The present text was reviewed by Facebook before publication for the purpose of identifying specific types of confidential information. Given the special nature of the embeddedness of the research we considered this appropriate and in keeping with research ethics. No removal of information was requested by Facebook.

¹ ECtHR, *Cengiz and Others v. Turkey*, judgment of 1 December 2015, § 49.

² Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7 March 2018, PP 2.

³ *Ibid.*, para. 1.1.3.

⁴ Cf. Kettemann/Tiedeke, “Back up: Can Users Sue Platforms to Reinstate Deleted Content? A Comparative Study of US and German Jurisprudence on “Must Carry””, GigaNet Symposium Research Paper (November 2019), http://leibniz-hbi.de/uploads/media/Publikationen/cms/media/xppiyy0_191011Kettemann_Tiedeke_Giganet-Back-up.pdf.

Courts in the US regularly reject claims by users who argue that norms exterior to the private order of a specific provider of online communication services control a conflict emerging within a private communication setting. Such conflicts about the reinstatement of deleted content or deleted profiles, which boil down to the question of who *ultimately* controls private communication spaces, is usually solved by US Courts in favour of private communication services. In *Johnson v. Twitter Inc.*, the California Superior Court refused to consider Twitter analogous to a “private shopping mall”.⁵ In *Prager University v. Google LLC*, the Northern California District Court (2018)⁶ refused to see YouTube as a state actor in accordance with the “public function”-test, arguing that providing a video sharing platform fulfils neither an exclusive nor a traditional function of the state. In most cases (*including Mezey v. Twitter Inc.*⁷, *Twitter Inc. v. The Superior Court ex rel Taylor*,⁸ *Williby v. Zuckerberg*,⁹ *Fyk v. Facebook Inc.*¹⁰, *Murphy v. Twitter, Inc.*¹¹ and *Brittain v. Twitter Inc.*¹²), US courts make use of the strong protection provided by Sec. 230 CDA, which largely immunizes intermediaries. To this day, no US court has managed to rule that an online communication platform should be liable to a “must-carry” claim.¹³

German courts, however, take a markedly different approach. With reference to Article 5 (1)(1) of the Basic Law (GG) and the judicially perceived function of influential providers of private communication spaces as a “public marketplace” (without, however, assuming state-like communication-enabling function¹⁴) a growing number of courts have decided, spearheaded by the introduction of the Network Enforcement Act,¹⁵ that Facebook and YouTube¹⁶ would generally¹⁷ not be allowed to remove “admissible expressions of opinion”, or at least would have to treat users equally in so doing, and that community standards would be interpreted in light of the horizontal effect of fundamental rights (a discussion led in German public law doctrine under

⁵ *Pruneyard Shopping Center v. Robins*, 9.6.1980, 447 U.S. 74.

⁶ *Prager University v. Google, LLC*, Northern California District Court, 2018 WL 1471939.

⁷ *Mezey v. Twitter Inc.*, Florida Southern District Court 1:18-CV-21069.

⁸ *Twitter Inc. v. The Superior Court ex rel Taylor* - A154973.

⁹ *Williby v. Zuckerberg*, Northern California District Court, 18-cv-06295-JD.

¹⁰ *Fyk v. Facebook Inc.*, Northern California District Court, No. C 18-05159 JSW.

¹¹ *Murphy v. Twitter Inc.*, San Francisco Superior Court, CGC-19-573712.

¹² *Brittain v. Twitter Inc.*, Northern California District Court, 2019 WL 2423375.

¹³ Keller, Daphne, “Who do you sue? State and platform Hybrid power over online speech”, Hoover Institution Essay, National Security, Technology and Law Aegis Series Paper No. 1902, 2.

¹⁴ Higher Regional Court Dresden (OLG Dresden), decision of 8 August 2018 – 4 W 577/18 at 19 et seq.; Higher Regional Court Karlsruhe (OLG Karlsruhe), decision of 28 February 2019 – 6 W 81/18 at 51 et seq.; Higher Regional Court Karlsruhe (OLG Karlsruhe), decision of 25 June 2018 – 15 W 86/18 at 21; Higher Regional Court Stuttgart (OLG Stuttgart), decision of 6 September 2018 – 4 W 63/18 at 71; Regional Court Offenburg (LG Offenburg), judgment of 20 March 2019 – 2 O 329/18 at 80; Regional Court Bremen (LG Bremen), judgment of 20 June 2019 – O 1618/18 at 59; Regional Court Heidelberg (LG Heidelberg), judgment of 28 August 2018 – 1 O 71/18 at 38.

¹⁵ See Schulz, Wolfgang / Kettemann, Matthias C. / Heldt, Amélie P. (2019): “Problems and Potentials of the NetzDG – five expert opinions”, (Hamburg: Verlag Hans-Bredow-Institut, 2019) (Working Papers of the HBI #48) (Arbeitspapiere des HBI Nr. 48).

¹⁶ Higher Regional Court Berlin (KG), decision of 22 March 2019 – 10 W 172/18 at 17.

¹⁷ Higher Regional Court Munich (OLG München), decision of 28 December 2018 – 18 W 1955/18 at 19 et seq. – possible exception for subforums.



the heading of *Drittwirkung*).¹⁸ However, the jurisprudence related to the Network Enforcement Act is far from consolidated.

A focus on the liability of intermediaries, however, glosses over a normative phenomenon that is unique to communication practices in the digital age: private actors set the rules for privately owned communicative spaces. These intermediaries, including social media companies, have become important normative actors,¹⁹ especially as the number of truly global ones is not big: Network effects and mergers have led to the domination of the market by a relatively small number of key companies. These companies have rights and obligations under international law and national legal systems. In line with the UN's Guiding Principles on Business and Human Rights and the "Protect, Respect and Remedy" Framework ("Ruggie Principles"),²⁰ intermediaries should respect the human rights of their users and other affected parties in all their actions (including the formulation and application of terms of service) and to provide redress for violations throughout their business relationships up and down the value chain. However, this leaves them with substantial freedom to normatively structure their platform through terms of service within the limits of the (up to) 193 national jurisdictions in which they are active.

Research on non-statal governance structures in the past mostly targeted multistakeholder-based governance structures and treated organizations as monoliths without considering the dynamics of internal norm-creation and norm-development processes.²¹ More recent literature has also focused on how social media companies apply rules: the moderation practices and the moderators involved, as well as automated content moderation systems in use have

¹⁸ Higher Regional Court Munich (OLG München), decision of 17 July 2018 – 18 W 858/18 at 30, ; decision of 12 December 2018 – 18 W 1873/18 at 21; decision of 17 September 2018 – 18 W 1383/18 at 20 et seq.; decision of 24 August 2018 – 18 W 1294/18 at 28; Regional Court Karlsruhe (LG Karlsruhe), decision of 12 June 2018 – 11 O 54/18 at 12; Regional Court Frankfurt/Main (LG Frankfurt/M.), decision of 14 May 2018 2-03 O 182/18 at 16; Regional Court Bamberg (LG Bamberg), judgment of 18 October 2018 – 2 O 248/18 at 86.

¹⁹ With 2.41 billion monthly active users (Q2/2019), Facebook is the biggest social network. See Statista, Number of monthly active Facebook users worldwide as of 2nd quarter 2019 (in millions), <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide>.

²⁰ See the "Ruggie Principles": Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie, Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework, UN Doc. A/HRC/17/31 of 21 March 2011.

²¹ Suzor, Nicolas, "Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms" at page 2 et seq.; Burris, Scott; Kempa, Michael and Shearing, Clifford, "Changes in governance: A cross-disciplinary review of current scholarship", (2008) Akron Law Review, 41, 1-66 (3); Black, Julia, "Decentering regulation: Understanding the role of regulation and self-regulation in a "post-regulatory" world", 2001 Current Legal Problems, 103-146 and Julia Black, "Constructing and contesting legitimacy and accountability in polycentric regulatory regimes", Regulation & Governance (2008) 2, 137-164; Parker, Christine "The pluralization of regulation", Theoretical Inquiries in Law, 349-369; Redeker, Dennis, Gill, Lex and Urs Gasser, "Towards digital constitutionalism? Mapping attempts to craft an Internet Bill of Rights", International Communication Gazette (2018), 302-319 (308).



been, up to now, extensively documented.²² But the main focus of studies in the field has hitherto erred on the side of norm execution and not norm-setting. More recently, however, there have been a number of scientific²³ investigations that look into the impact of private norm-setting and selected journalistic reports²⁴ on private norm-making at Facebook. There have, however, not yet been any studies that have systematically and empirically investigated the complex processes through which social networks develop norms, who they involve in these processes and – importantly – how actors within these companies conceive of their norm-setting function. This can be easily explained by the difficulty of conducting empirical research within these companies. Regarding Facebook, Inc., one author found a lack of “amenability to demystify and open the black box of content guideline creation”. He called it “the one area where Facebook is not actually willing to make concessions.”²⁵

²² See Tarleton Gillespie, “Custodians of the Internet” (pp.116-135), where he discusses “the human labor of moderation” and distinguishes between the work of “internal teams, crowdworkers, community managers, flaggers, superflaggers, peer support and external efforts and everyone”.

²³ See Thomas E. Kadri and Kate Klonick, “Facebook v. Sullivan: Public Figures and Newsworthiness in online speech” (p. 22): “The lack of transparency and accountability gives little comfort to those who worry about the mercurial and subjective nature of newsworthiness determinations at Facebook”, available at <https://knight-columbia.org/content/facebook-v-sullivan>; Klonick, Kate; “The New Governors: The People, Rules, and Processes Governing Online Speech”, (2018) 131 Harv. L. Rev. (1598) at page 1657 noting that: “As new situations arise during moderation, platforms will both tweak current policy as well as develop new rules. Many of these judgments continue to be difficult to make.” Suzor, Nicolas, “Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms”, at 4, arguing that the rule of law “can be usefully applied to assess the governance of digital media, paying particular attention to the role of platforms as writers of the rules of participation; designers of technology that enables communication and constrains action; developers of algorithms that sort, organize, highlight, and suppress content; and employers of human moderators who enforce rules on acceptable content and behavior”; Suzor, Nicolas, “Lawless – the secret rules that govern our digital lives”, at 114, warns that “[T]he central challenge for internet governance is now to find new ways to ensure that the power of online intermediaries over our lives is exercised in a way that is fair and accountable, without destroying the massive benefits that an open and diverse global internet can bring”; Gillespie, Tarleton, “Custodians of the Internet”, at 5 explains that “[...] whether they want to or not, platforms find that they must serve as setters of norms, interpreters of laws, arbiters of taste, adjudicators of disputes, and enforcers of whatever rules they choose to establish. Having in many ways taken custody of the web, they now find themselves its custodians.”

²⁴ Simon van Zuylen-Wood, Vanity Fair, March 2019, “Men are scum”: inside Facebook’s war on hate speech, <https://www.vanityfair.com/news/2019/02/men-are-scum-inside-facebook-war-on-hate-speech>: “for a user base of more than two billion people, such changes proved impossible to scale. On some level, there are no “fixes” to Facebook’s problems. There are only trade-offs. Like an actual government, it seemed, the best Facebook could hope for was a bunch of half-decent compromises. And like a government, anything it did would still piss off at least half its constituents” quoting Monika Bickert, who concedes that “[T]he world is too diverse,” ... “And people see speech so differently, and safety so differently. I don’t think we’re ever going to craft the perfect set of policies where we’re like, ‘We’ve nailed it.’ I don’t think we ever will”; Jason Koebler and Joseph Cox, “Facebook’s Struggle to Moderate Two Billion People”, Vice, 23 August 2018, describing content moderation as “one of the most labor-intensive and mind-bogglingly complex logistical problems Facebook has ever tried to solve [...] Facebook believes highly-nuanced content moderation can resolve this tension, but it’s an unfathomably complex logistical problem that has no obvious solution, that fundamentally threatens Facebook’s business, and that has largely shifted the role of free speech arbitration from governments to a private platform.”

²⁵ David Morar, Facebook’s Oversight Board: A toothless Supreme Court?, 2 October 2019, Internet Governance Project, <https://www.internetgovernance.org/2019/10/02/facebooks-oversight-board-a-judiciary-with-no-constitution>.



In the context of the present project we were given access to this “black box” of normative development within the social network. In addition to many background interviews with employees responsible for different aspects of the crafting of new policies, a researcher from this project was able to observe in person the actors involved and their relationships within Facebook Inc. We are, therefore, in a position to map Facebook Inc.’s field of private development in regard to content-related policies and formulate initial insights into some of the challenges posed by private norm-setting through the example of one specific social network.

This research paper is the first based on research conducted as part of this pilot study. Section two describes our methodological approach while section three goes on to present Facebook’s own description of the process they go through when adopting new norms. Section four will then contrast this description with an analysis of the data gathered through our research. Section five aims to wrap up the comparison of our participants’ observations and our analysis in a comprehensive conclusion.



2. METHODOLOGICAL APPROACH

Our research is essentially an empirical case study carried out in the tradition of social science and anchored in regulation studies (*Regelungs- und Regulierungswissenschaft*). Our research is embedded within a theoretical framework that adopts Norbert Elias' figuration theory²⁶ to help us better understand how social reality is never a *given* but is only ever produced, reproduced, criticized, and altered through the interrelations between actors in selected settings, that is, within what Elias labels as "figurations". This is also true for the production of rules in a specific societal domain.²⁷ Approaching (internet) governance – the creation and application of norms, rules, and standards that impact the use and development of the internet – this analytical framework provides certain routes through which we are able to pin down our methodology: First, we are able to conduct hermeneutic analyses of how normative choices are realized in and through contracts, terms of service, laws and code and, thus, reveal (hidden) structures. Second, we can observe the figurations involved and analyse their particular features and practices, for example, by interviewing the actors involved, tracking and analyzing discourses or by observing those figurations through a set of digital ethnographies.

As a domain-specific concept analyzing the normative dimension of the standard-setting process, based on the empirical research and with a view to the specific *situatedness* (member, shared intentionality, practices) of the observed communicative figuration, we employ the *normative orders* approach as developed by Rainer Forst and Klaus Günther and the *Frankfurt School of Critical Normative Research*.²⁸ A normative order is a "complex of norms and values with which the fundamental structure of a society [or a sub-section of it](...) is legitimated, in particular the exercise of political authority and the distribution of basic goods."²⁹

For the purposes of this study, the normative order will be considered as the observed complex of norms, values and practices – established through our empirical research and situated through the *communicative figurations* approach – that relate to, shape, and legitimize the development (adaption, responsiveness, contestation) and use of community standards.

²⁶ Elias, Norbert, "What is sociology?" (London: Hutchinson, 1978).

²⁷ Cf. Ziewitz, Malte and Pentzold, Christian, "In search of internet governance: performing order in digitally networked environments", (2014) *New Media & Society*, 16(2): 306-322.

²⁸ Frankfurt School of Critical Normativity around Rainer Forst and Klaus Günther at the Cluster of Excellence *The Formation of Normative Orders*. See Rainer Forst and Klaus Günther, "Die Herausbildung normativer Ordnungen. Zur Idee eines interdisziplinären Forschungsprogramms", in Rainer Forst and Klaus Günther (eds.), "Die Herausbildung normativer Ordnungen. Interdisziplinäre Perspektiven", (Frankfurt/New York: Campus, 2011), 11-30 (15) (our translation). See Matthias C. Kettmann (eds.), *Normative Orders. Fresh Perspectives* (Frankfurt/New York: Campus, 2020).

²⁹ Forst and Günther (2011).



In order to operationalize the figuration concept and apply it to the private ordering of internet-based communication we use Uwe Hasebrink and Andreas Hepp's communication-based development of Elias' figurations, the "communicative figurations".³⁰ These communicative figurations have three distinctive elements:

- (1) a communicative figuration maintains a certain *constellation of actors* that can be regarded as its structural basis, a network of individuals who are interrelated and are communicating;
- (2) each communicative figuration has at its core dominant *frames of relevance* that serve to guide its constituting practices and form a shared identified, or inferable, intentionality;
- (3) figurations are based on specific *communicative practices* that are interwoven with other social practices³¹ and which give the figurations meaning.³²

Based on these theoretical considerations our key question is how the "Community Standard" (and other relevant norms) for the social media platform Facebook are produced, by which constellation of actors, under what frame of relevance, and in using which practices.

During the preparation stage of this study we received feedback on our methods by researchers from disciplines ranging from law through the sociology of media, to ethnography. The empirical part is multi-methodologically designed and combines participatory organizational observation influenced by ethnography (August 2019),³³ expert interviews and document reviews (July 2019 and September 2019). One co-author participated in meetings with Facebook's Product Policy team for a one week-period in late August 2019. Before this period a research guide was drafted to objectify observational approaches and control for disciplinary bias. We opted for a detailed research guide since the observer was an expert in internet law and the guideline was meant to ensure focus on the observation and not to prematurely infer normativity in the collected data.

Within our ethnographically-informed participatory organizational observation³⁴ we employed the "participant as observer" stance, that is, the researcher was present in the group being studied, and the group was aware of the research activity, but the researcher did not take an

³⁰ Hepp, Andreas, Breiter, Andreas and Hasebrink, Uwe, "Rethinking Transforming Communications: An Introduction", in Andreas Hepp, Andreas, Breiter, Andreas and Hasebrink, Uwe, "Communicative Figurations. Transforming Communications in Times of Deep Mediatization", London: Palgrave, 2018), 3-15 (7).

³¹ Hepp, Andreas and Hasebrink, Uwe "Researching Transforming Communications in Times of Deep Mediatization: A Figurational Approach", in Andreas Hepp, Andreas, Breiter, Andreas and Hasebrink, Uwe, "Communicative Figurations. Transforming Communications in Times of Deep Mediatization", (London: Palgrave, 2018), 15-50.

³² Hepp, Andreas, Breiter, Andreas and Hasebrink, Uwe "Rethinking Transforming Communications: An Introduction", in Hepp et al. (eds.), (2018), 3-15 (7).

³³ Over the course of one week in late August 2019 and via the internet in October and November 2019 the project researcher participated in working group meetings and interviews. In all, we have made 751 numbered observational notes. We have stored the document containing all the researcher's notes and those with other research materials in the HBI Science Data Repository but cannot provide access for replication studies due to our agreement with Facebook.

³⁴ Cf. Ziewitz, Malte and Pentzold, Christian "In search of internet governance: performing order in digitally networked environments", (2014) *New Media & Society*, 16(2): 306-322.





active part in the activity.³⁵ The observation process was structured to maximize the efficiency of the “field” experience.³⁶ This presupposed agreement between the researcher and the participants as to their presence and role.

The goal of the case study was to plot, situate and analyze the development of content-related policies within an important social media service provider to reconstruct how private norms that impact public communication are developed, how (and why) the input of external actors is strategically sought, whether or not the legitimacy of these norms is considered, and if so, how this is proceduralized, and the self-perception of the actors involved. Our research tracked the cycle of policy change through fifteen parallel processes. The earliest of these still active at the start of our project had begun in early summer 2019. We consciously selected the second half of the year 2019 as our research phase in light of the changes the Oversight Board that will be established in January 2020 will most likely bring to the table.

Additional expert interviews and online participation in processes of normative change took place throughout October and November 2019. The data were then analysed through *deep and dense description* processes (October 2019).³⁷ Final interviews with key participants that allowed for critical reflection of the immersive approach took place in early December 2019. This data-grounded, normative-epistemic reflection, systemization (and critical narrativization) of disparate collected data points and their analytical rationalizations allow us to reconstruct underlying (organizationally deeply situated) patterns of sense and meaning.

The present paper is the preliminary result of a deep and dense reflection on the data collected in the first cycle of research, its figurational ordering and normative analysis. Being based on a case study the results merely reflect the processes of rule-changing under analysis and do not allow for generalization. We offer just one glimpse, at a specific point in time, within a long process of developing a normative order of communication. It is a snapshot.

³⁵ Gold, R.L. “Roles in sociological field observations”, (1958), *Social Forces*, 36, 217-223.

³⁶ Angrosino, M. V. and Mays de Perez, K. A. , “Rethinking observation: From method to context”, 2000 in: Norman K. Denzin & Yvonna S. Lincoln (eds.), *Handbook of Qualitative Research* (2nd ed.) pp. 673-702, Thousand Oaks, CA: Sage.

³⁷ Cf. Geertz, Clifford, “Thick Description: Towards an Interpretive Theory of Culture”, in Clifford Geertz (ed.), *The Interpretation of Cultures*, Basic Books (1973).





3. CONCEPTUALIZING GOVERNANCE: HOW FACEBOOK APPROACHES THE POLICY DEVELOPMENT PROCESS

To understand the results of the pilot study it is important to grasp the modalities of normative development of the Community Standards as conceptualized by Facebook. In this section we therefore present – based on (mostly) publicly available self-descriptive documents with our understanding of them influenced by internal discussions – which *constellations* (“*figurations*”) of actors are involved in processes of normative change (3.1.), what norms, standards, and commitments frame it (*shared frames of relevance*) (3.2.) and what the multi-part ideal policy change process looks like (*proceduralized communicative practices*) (3.3.).

3.1. Constellations of Actors

What constellations (“figurations”) of actors are involved? The Product Policy team is in charge of the process of adapting Community Standards to present needs. Key players during the policy change process are the thematic coordinators of the team – the “subject-matter experts (SMEs)” –, the internal experts with whom they consult, the heads of department within Facebook involved in the policy development process, the Director for Human Rights, and the Stakeholder Engagement team that coordinates interaction with external stakeholders.

How do they interact? Before the first meeting of any working group the coordinator of the policy change, in accordance with the head of the Stakeholder Engagement team, develops an outreach strategy which is presented and agreed upon by the working group. Between the Working Group’s first meeting and subsequent meetings, outreach is conducted and plotted out on an overview (and later a slide for presentations to internal stakeholders) showing (any) diversity of opinions. Stakeholder comments on suggested changes are sought from a variety of non-governmental and academic sources. This is motivated by the expressed goal of the Product Policy team at Facebook to base its policies “on feedback from community representatives and a broad spectrum of the people who use our service”. The advice of experts is also sought. The process is meant to “root [Facebook’s] policies in sources of knowledge and experience that go beyond Facebook”.³⁸

Stakeholders are defined as “organizations and individuals who are impacted by, and therefore have a stake in, Facebook’s Community Standards. [...] our more than 2.7 billion users are, in the broad sense, stakeholders.”³⁹ Since consulting a third of the world’s population is difficult,

³⁸ Facebook, Stakeholder Engagement, https://www.facebook.com/communitystandards/stakeholder_engagement.

³⁹ Ibid.



and previous attempts to have users vote on standards have failed,⁴⁰ Facebook uses a proxy-oriented representative approach, “think[ing] of stakeholders as those who are informed about and able to speak on behalf of others.”⁴¹ The “primary focus of [its] engagement” are “civil society organizations, activist groups, and thought leaders, in such areas as digital and civil rights, anti-discrimination, free speech, and human rights.”

The Stakeholder Engagement team, a part of the Product Policy team, “also engage[s] with academics who have relevant expertise.” While these may not “directly represent the interests of others, [...] they are important stakeholders by virtue of their extensive knowledge, which helps [Facebook] create better policies for everyone.”⁴²

In conducting stakeholder engagement the team commits – in their own view – to the principles of inclusiveness (“to deepen our local knowledge and perspective – [to] hear voices we might otherwise miss”), expertise (“we don’t have all the answers”) and transparency (“opening up [the] policy-making process helps build trust”).⁴³

Before the second working group meeting on a given product policy change, the advice of Facebook’s internal academic research team is sought. Its responsibility is to distil key research insights from a wide variety of sources and present them concisely. At the second working group meeting summaries of external feedback are presented and the slide for the Product Policy Forum is prepared. At the Product Policy Forum a “detailed summary” of the feedback is presented. In the Forum’s minutes⁴⁴, the “range and nature of engagement” is reflected as it is the “rationale for [Facebook’s] final decision”.⁴⁵ At a later stage, stakeholders are informed about the impact they have had on the process.

3.2. Shared Frames of Relevance

Which norms, norm systems, values and commitments are in play during the policy development process? The detailed set of Community Standards that is clearly necessary today has only emerged over recent years. In her review of the evolution of the moderation practices with Facebook, Kate Klonick quotes Facebook’s Dave Willner on the status quo in 2009 as follows:

⁴⁰ Cf. Suzor: Nicolas, “Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms”, *Social Media and Society* (2018), Volume: 4 issue: 3, <https://journals.sagepub.com/doi/10.1177/2056305118787812>.

⁴¹ Facebook, Stakeholder Engagement, https://www.facebook.com/communitystandards/stakeholder_engagement.

⁴² Ibid.

⁴³ Ibid.

⁴⁴ Cf. Facebook, Product Policy Forum (formerly Content Standards Forum), Minutes, <https://newsroom.fb.com/news/2018/11/content-standards-forum-minutes>.

⁴⁵ Facebook, Stakeholder Engagement, https://www.facebook.com/communitystandards/stakeholder_engagement.



"The [policy] guidance was about a page; a list of things you should delete: so it was things like Hitler and naked people. None of those things were wrong, but there was no explicit framework for why those things were on the list."⁴⁶

A former Facebook employee is quoted as describing the guiding content moderation rule in 2008 as follows: "if it makes you feel bad in your gut, then go ahead and take it down."⁴⁷ Since then, Facebook's approach to content moderation based on defined Community Standards has evolved substantially. In late 2009, Jud Hoffmann was hired as Global Policy Manager and a team was charged with developing publishable "Community Standards"⁴⁸ of no more than 15,000 words.⁴⁹ The guiding principle, as expressed in the 2009 iteration of the Community Standards, was to give "people the power to share and make the world more open and connected."⁵⁰ The statement did not yet invoke "voice" or refer to, or be premised upon, a specific vision of a more open and connected world. From 2010 onwards, Facebook has gradually adapted its Community Standards and greatly increased their granularity.

It was, however, not until 2017 that the founder and CEO of Facebook, Inc., Mark Zuckerberg, revisited the mission statement and acknowledged that Facebook had made mistakes in the past and announced it would "develop the social infrastructure to give people the power to build a global community that works for all of us."⁵¹ Facebook's new goal was conceptualized as giving "people the power to build community and bring the world closer together."⁵² In 2018, internal training materials for content moderators were leaked.⁵³ Later that year, Facebook decided to publish more details of how their Community Standards were enforced.⁵⁴

In 2019, two changes to the normative framework in which Facebook operated were implemented. Monika Bickert, Vice President of Global Policy Management and chair of the Product Policy Forum, in which all policy changes are discussed, said that their main goal was less introducing new rules but rather providing "clarity".⁵⁵ The first change was a stronger enunciation

⁴⁶ Klonick, Kate, "The New Governors: The People, Rules, and Processes Governing Online Speech", (2018) 131 Harv. L. Rev. (1598) at 1631.

⁴⁷ Ibid..

⁴⁸ Ibid., 1620.

⁴⁹ Ibid.

⁵⁰ History of Mission Statements: Gillian Reagan, "The evolution of Facebook's Mission Statement", 13 July 2009 <https://observer.com/2009/07/the-evolution-of-facebooks-mission-statement>.

⁵¹ Mark Zuckerberg, Building Global Community, 16 February 2017, <https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634>.

⁵² Mark Zuckerberg, Video live from the Facebook Community Summit in Chicago, 22 June 2017, <https://www.facebook.com/zuck/videos/10103817960742861>.

⁵³ Max Fisher, "Inside Facebook's Secret Rulebook", New York Times, 27 December 2018, <https://www.ny-times.com/2018/12/27/world/facebook-moderators.html>.

⁵⁴ Facebook, Data Snapshot: Facebook's Community Standards Enforcement Report October 2017-March 2019, <https://fbnewsroomus.files.wordpress.com/2019/05/cser-data-snapshot-052219-final-hires.png>.

⁵⁵ Facebook, Community Standards, <https://newsroom.fb.com/news/2018/04/comprehensive-community-standards>.



of certain values in the preamble of the Community Standards in an update⁵⁶ and the second was the announcement of the creation of an Oversight Board and its respective Charter.⁵⁷

Facebook's overarching goal – according to its own communications – is to give people “voice”: “Our commitment to giving people voice remains paramount.” As a result, there is a preference against deletion. Even content violating the Community Standards can stay online, if the “voice”-dimension is more important in light of a newsworthiness and public interest test. Such an assessment is conducted “after weighing the public interest value against the risk of harm” in light of “international human rights standards”.⁵⁸

Expressions can be limited in order to pursue any or all of the following goals: authenticity, safety, privacy, and dignity.⁵⁹ In light of these values, Facebook professes to not wanting “people using Facebook to misrepresent who they are or what they're doing”, trying to minimize expression “that threatens people [and] has the potential to intimidate, exclude or silence others”, with the aim of protecting “personal privacy and information” and ensuring that users “respect the dignity of others and not harass or degrade others”.⁶⁰

Voice is enhanced by notions of newsworthiness and public interest. These interests are balanced against the risk of harm and limited in the interest of authenticity, safety, privacy, and dignity. Taken together these are the values pursued in the normative process of product policy development. While the Community Standards make a reference to international human rights, they do so only in referring to identifying an exception for content that should stay online even if it violates Community Standards⁶¹ (human rights standards are also referred to the Charter of the Oversight Board.⁶²).

⁵⁶ Monika Bickert, Vice President, Global Policy Management, Facebook, Updating the Values That Inform Our Community Standards, 12 September 2019, <https://about.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards>.

⁵⁷ Facebook, Structure of the Oversight Board, <https://newsroom.fb.com/news/2019/09/oversight-board-structure>.

⁵⁸ Ibid.

⁵⁹ Ibid.

⁶⁰ Monika Bickert, Vice President, Global Policy Management, Facebook, 12 September 2019, “Updating the Values That Inform Our Community Standards”, <https://newsroom.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards>.

⁶¹ Facebook, Community Standards, <https://www.facebook.com/communitystandards>. One example is the controversy generated around the “Napalm Girl” article and photos: See Sam Levin, Julia Carrie Wong, and Luke Harding, The Guardian, “Facebook backs down from ‘napalm girl’ censorship and reinstates photo”, 9 September 2016.

⁶² Introduction of the Charter: “Freedom of expression is a fundamental human right”, Article 2 Sec. 2 Sentence 3 Charter: “When reviewing decisions, the board will pay particular attention to the impact of removing content in the light of human rights norms protecting free expression”. https://fbnewsroomus.files.wordpress.com/2019/09/oversight_board_charter.pdf



Apart from its Terms of Service⁶³ Facebook has established Community Standards as its main set of norms.⁶⁴ In six different parts of the Community Standards Facebook refers to its values.⁶⁵ The Community Standards also provide for definitions regarding certain terms that restrict voice or content.⁶⁶ The introduction starts with a commitment to voice: “To ensure that everyone’s voice is valued, we take great care to craft policies that are inclusive of different views and beliefs, in particular those of people and communities that might otherwise be overlooked or marginalized.”⁶⁷ Restrictions that are made on content in the context of potential offline harm or personal safety are explained in Sec. I of the Community Standards. Facebook, however, pledges to make the best efforts to consider context and language when deciding to restrict voice.⁶⁸ Sec. III provides details relating to the standards set out by Facebook and on objectionable content including hate speech, violence and graphic content, adult nudity and sexual activity, sexual solicitation, and cruel and insensitive content. Sec. IV explains the dimensions of integrity and authenticity along the lines of the convictions laid out in the introduction, “that authenticity creates a better environment for sharing, and that’s why we don’t want people using Facebook to misrepresent who they are or what they’re doing.”⁶⁹ Sec. V refers to the safeguard of IP rights and Sec. VI details the procedure employed in cases of user requests.

Facebook’s expressed values have corresponding rights in international human rights law. Voice is protected by freedom of expression, (one of) the most fundamental human rights, which is protected by international, regional, and national human rights law. Apart from the fact that voice shall be the paramount principle, the values have no clearly visible hierarchy. Facebook states that it will establish their relationship by balancing voice against the risk of harm, while also taking into account international human rights standards related to speech and expression.⁷⁰

Privacy and dignity are constitutional values that are explicitly protected in all liberal democracies and by the International Bill of Rights and regional human rights conventions.⁷¹ The Community Standards do not explicitly refer to these documents, rather they state that Facebook would “look at international human rights standards to make these judgments.” The same

⁶³ Facebook Terms of Service, <https://www.facebook.com/legal/terms>.

⁶⁴ Other terms and policies include: Commercial Terms, Advertising Policy, Self-Serve Ad Terms, Pages, Groups and Events policy, Facebook Platform Policy, Developer Payment Terms, Community Payment Terms, Commerce Policies, Facebook Brand Resources and Music Guidelines, <https://www.facebook.com/legal/terms>.

⁶⁵ Voice: Introduction; Safety: Sec. II no. 7, 9, 10, 12 Sec. IV no. 19.; Authenticity: Introduction, Sec. III, no. 13 Sec. IV no. 18, 19, 20, 23; dignity(dehumanizing): Introduction; Sec. III no. 23; Privacy, Introduction, Sec. II no. 12. <https://www.facebook.com/communitystandards/introduction>.

⁶⁶ Facebook, Community Standards, Introduction, <https://www.facebook.com/communitystandards/Introduction>.

⁶⁷ Ibid.

⁶⁸ Community Standards Sec. I Nr. 1 Violence and Incitement.

⁶⁹ Facebook, Community Standards, Introduction, <https://www.facebook.com/communitystandards/introduction>.

⁷⁰ Facebook, Community Standards, <https://www.facebook.com/communitystandards>.

⁷¹ Dignity: Article 1 UDHR, Article 5 ACHPR (Banjul Charter), Article 5 no. 2 and Article 11 no. 1 IACHR and Article 10 no. 1 ICCPR Privacy: Article 12 UDHR, Article 17 ICCPR, Article 11 IACHR. The ACHPR (Banjul Charter) does not specifically refer to privacy.



is true for the described method of “weighing the public interest value against the risk of harm”. This is important because a further value, if not at par with the others, is the public interest character of content, basically its “newsworthiness”. While content that violates Community Standards is usually deleted, in exceptional cases a wider range of expression is allowed in the interest of public debate.⁷²

Facebook commits to consistently and fairly applying its policies:

The consequences for violating our Community Standards vary depending on the severity of the violation and the person's history on the platform. For instance, we may warn someone for a first violation, but if they continue to violate our policies, we may restrict their ability to post on Facebook or disable their profile. We also may notify law enforcement when we believe there is a genuine risk of physical harm or a direct threat to public safety.⁷³

Although the Community Standards do not explicitly reference proportionality, the method described (“vary depending on the severity [...] and the person's history”) invokes some elements of a traditional proportionality test.⁷⁴

This normative order of Facebook has changed with the publication of the Charter of the Oversight Board⁷⁵ and the content governance structure it heads.⁷⁶ The Charter opens with a commitment to freedom of expression as a fundamental human right.⁷⁷ In its introduction, it restates that free expression remains the paramount principle but may be limited in order to facilitate someone else's voice or when it is in conflict with the four other principles: authenticity, safety, privacy and dignity.

With the introduction of the Charter, Facebook accepts the responsibility to set standards which are based on values and the responsibility to ensure “fair decision-making” to operate transparently and to “articulate” and “explain” these steps to the public.⁷⁸

Throughout late 2019 Facebook's Public Policy team conducted work on another normatively relevant frame. The *Policy Prioritization Framework*, which was still tested at the time of publication of the present paper, was designed to better gauge which policy changes were to be

⁷² New York Times Co. v. Sullivan, 376 U.S. 245 (1964). See Kadri and Klonick.

⁷³ Facebook, Community Standards, <https://www.facebook.com/communitystandards>.

⁷⁴ Sweet, Alec Stone, and Jud, Mathews, Proportionality Balancing and Global Constitutionalism, Columbia Journal of Transnational Law, vol. 47, no. 1, 2008, 72-164 (74 et seq.).

⁷⁵ Nick Clegg, the Vice President of Global Affairs and Communications at Facebook announced that the “Oversight Board will make Facebook more accountable and will improve (...) decision-making.” <https://newsroom.fb.com/news/2019/09/oversight-board-structure>.

⁷⁶ Brent Harris, Director of Governance and Global Affairs, “Establishing Structure and Governance for an Independent Oversight Board”, 17 September 2017, <https://newsroom.fb.com/news/2019/09/oversight-board-structure>. While the Charter of the Oversight Board was only published after the research conducted into the Product Policy team, the coming changes were already influencing discussions and approaches, so it is fair to include the Charter as a normative tool in our analysis of the frame of reference.

⁷⁷ First sentence of the introduction: “Freedom of expression is a fundamental human right”. Again, in Article 2 Sec. 2 Sentence 3: “When reviewing decisions, the board will pay particular attention to the impact of removing content in light of human rights norms protecting free expression.” Facebook, Oversight Board Charter, https://fbnewsroomus.files.wordpress.com/2019/09/oversight_board_charter.pdf.

⁷⁸ Ibid.



prioritized over others, and through which normative channel they were to be routed. Two leading questions lead the prioritization efforts: (1) "Change": "What would be the impact of change?" measured with the metrics of a *Severity and Harm Index*, the problem's prevalence, and the operational feasibility of the changes, i.e. that its implementation would not lead to cognitive overload for moderators. (2) "Contestability" (with perhaps "contentiousness" better expressing the normatively relevant character of so qualifying a change: "Is the topic under review rather settled globally and within Facebook or are there strong and contentious views and laws on key aspects?", measured with the metrics of "tension between principles", "lack of consensus", and any previously received "leadership input". Policy proposals that score high on the change and contestability metrics – (important) and high (contentious) or low (non-contentious) – would usually be routed through normal policy development channels, including the Public Policy Forum. Policy proposals scoring low on importance/change and low on the non-contentious normative setting/contestability metric would be pursued, but with less emphasis on speed. Policy proposals scoring low on importance and high on contentiousness are pursued in different ways, including updates to operational guidance and training materials.

3.3. Patterns of Communicative Practices

Any policy development (rule change) in the community standards follows a defined multi-step process. First, Facebook's product policy department collects *signals* that may come from societal trends, the content review teams, feedback from quality assessment, and internal or external research. As these signals are collected, some of them solidify into a priority for policy development. During this phase internal stakeholders are alerted to possible changes. At a head-up/kick-off meeting a policy development is initially discussed within the Product Policy team. Then the actual policy *development* process begins with the topic assigned to a coordinator who introduces a potential change in community standards at the Product Policy Forum, a meeting of about thirty employees at Facebook Headquarters (HQ) and some 75 from offices across the world representing, inter alia, teams from Policy, Community Operations, Community Integrity and Communications.⁷⁹

Subsequent to this introductory phase, a working group meeting takes place to plan the necessary research on the issue behind the possible policy change and the stakeholder engagement with stakeholders from civil society and academia. The research department conducts a literature review and provides scientific expertise. A number of working group meetings are held in which internal stakeholders are also consulted. At this stage feedback from Facebook's upper management team, the "leadership", is sought and incorporated. After a moot presentation of the change before the working group, the proposals are finalized and policy recommendations are then formulated.

At a later Product Policy Forum the development in policy is introduced, alternatives are weighed up, and a new community standard (or a change to an existing one) is adopted (or not).

⁷⁹ As observed during participation in a number of Product Policy Forum sessions.





Facebook publishes the minutes of the Product Policy Forum, including slide decks, online. At the time of writing the slide decks from 13 November 2018 to 7 January 2020 were online. There is usually a delay to ensure that no user data is accidentally published.⁸⁰

During the policy development process the Product Policy team also considers the implication of a change on moderators. During the implementation phase the language of the policy is finalized (sometimes in parallel to the change being tested), training materials are created, review tools are developed, and content review teams are trained. At the launch of the change, the policy change “goes live” and the Community Standards are formally updated.

⁸⁰ Cf. Facebook, Product Policy Forum (formerly Content Standards Forum), Minutes, <https://news-room.fb.com/news/2018/11/content-standards-forum-minutes>.



4. DOING GOVERNANCE: CASE STUDY-BASED ANALYSES

The regulatory structures and practices related to the development of product policy with Facebook are intricate. This section will demonstrate the results of our observations in contrast to Facebook's own conception of the normative process described above (Section 3) in light of the analytical triad: a description of the observed constellations of actors (4.1.), a reconstruction of the underlying frames of reference (4.2.) and an analysis of the communicative practices shaping the norms (and their production processes) (4.3.).

4.1. Constellation of Actors

One key observation is that external stakeholders are extensively referred to in the discourse on policy changes; from the "signals" that trigger the process through to the evaluation of the involved interests to the assessment of the impact of the policy put into place and the potential changes to policy in regard to those interests.

Stakeholders are selected based on the question, as one person we talked to put it, "Who do we think we should talk to in order to develop the best solution?"⁸¹ and conversations tended to go on until those responsible for stakeholder engagement felt convinced that they had heard enough *voices* to distil what normative outcome would work best to ensure *voice*. Meetings we attended usually ran the length of fixed time slots. However, on a number of occasions, when the issue at hand had not yet been discussed with enough clarity and to the satisfaction of the Facebook employee organizing the discussion, a further meeting was scheduled or a further talk planned.

The video-conferences we were able to attend all went very well and were characterized by a sincere interest in exchanging views on policy development, mutual appreciation, and an open exchange of ideas that contributed positively to the policy development process.⁸² The NGOs and academics on the other end of the line were eager to contribute and seemed to enjoy their experience. They seemed earnest, if sometimes at a loss about the granularity of a new norm, which, at least in one observed case, seemed to be beyond the comprehension of at least certain non-specialized NGOs. (Product Policy team members expressed awareness of this: "Sometimes the challenges we face are novel even to the experts we consult with."⁸³)

The breadth of engagements is represented with continuous references to external stakeholders — through different phases of the normative process — who are seen as integral parts of

⁸¹ Observation # 474, documentation of the embedded phase.

⁸² Observations # 223, 230 335, documentation of the embedded phase.

⁸³ Ibid.



change in policy. For example, stakeholders would be referred to in a manner such as: “a lot of the constitutional lawyers we talk to [...]”⁸⁴, or “Whenever we talk to experts they say [...]”.⁸⁵

The opinions of academics are often considered less biased even though they have adopted – if they have arrived at a certain station in their academic careers – certain rules and expectations.⁸⁶ We were also told of a number of workshops (“policy round-tables”) with academics on particularly difficult topics, primarily on topics that are still unregulated or philosophical questions. This was described to us as a new approach that had produced very positive results – the first one being a “proof of concept”, which “we’ll definitely do again”.⁸⁷

4.2. Shared Framing as Normative Work

In our observation the normative processes within the Product Policy team are based on the understanding of creating a separate and unique private normative order for public communication, including regulating “novel” questions, which aims at giving *voice* to users but is not directly linked to any national or international legal order. That is noteworthy since it is a *Product Policy* team and one might assume that their shared understanding is one of designing a service or a product.

Though we did not observe clear references to Facebook’s “constitutional” values during working group meetings, it is important to understand the framework within which the Product Policy team members act (or within which they believe that they act). When asking about the fundamental norms guiding the processes team members participated in, we were referred to the values expressed in the Community Standards and especially the principles developed together with an outside consultancy in business ethics (*Business for Social Responsibility*), as guiding Principles for the Stakeholder Engagement process.⁸⁸ Though the Community Standards have only been formally updated during the period that our project took place (“expanding the values”⁸⁹), we would tend to agree with the Product Policy team members who saw this move as not to represent a fundamental change, but rather, as a normative reflection of lived normativity.

We have had some difficulty identifying clear data points from which to gather insights into the lived relevance of the normative frames of relevance. Meetings on product policy were driven

⁸⁴ Observation # 474, documentation of the embedded phase.

⁸⁵ Observation # 586, documentation of the embedded phase.

⁸⁶ Rathbun, Gail and Turner, Nancy, “Authenticity in academic development: the myth of neutrality”, *International Journal for Academic Development*, 17:3, 231-242, DOI: 10.1080/1360144X.2012.679273 at 232.

⁸⁷ Observations # 260, 414, documentation of the embedded phase.

⁸⁸ Observations # 441-451, documentation of the embedded phase.

⁸⁹ “[W]e’re expanding the values that serve as the basis for our Community Standards” (Monika Bickert, Vice President, Global Policy Management, Facebook, 12 September 2019, Updating the Values That Inform Our Community Standards, <https://newsroom.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards>).



by the unstated yet cognizable goal of drafting a “good” standard that “makes sense”, and one that is “thoughtful”.⁹⁰ Trade-offs between “voice” and authenticity, safety, privacy, and dignity seemed to be internalized in the processes or were at least not formally discussed or visibly brought forward as issues for experts to engage with.⁹¹

While frames of relevance as unstated but influential spheres of meaning have a “compliance” pull, or a conforming function, it was difficult to find examples where their normative frame was made visible. Employees we talked to after our observation period pointed to Mark Zuckerberg’s 17 October 2019 talk at Georgetown on freedom of speech and marked it as “important” as it had clearly been previously discussed and shared internally. This would indicate that a certain normativity was structurally imbued within the principles developed there.⁹²

While Mark Zuckerberg referred to the mutual goals of “giv[ing] people voice, and bring[ing] people together” and described them as “go[ing] hand in hand”,⁹³ in our observations during team meetings voice is treated as a more powerful normative principle while inclusion (“bring people together”) tended to maintain a background role. The title of Mark Zuckerberg’s speech might point to this observation: “Standing for Voice and Free Expression”⁹⁴ does not immediately make apparent the importance of inclusion.

The stated frames of reference are enriched in our observation by a shared understanding that the company is creating a new normative order for more than two billion users, which had never been attempted before. Voice is, therefore, actually considered as an important normative lens that influences discussions: “We do not want to punish people” (for holding or voicing opinions), we were told.⁹⁵

A further frame of reference, not visible “on paper” but observable during the research process, was the conviction that the Product Policy team members were “doing good work”.⁹⁶ It was apparent that the Stakeholder Engagement team in particular was acutely aware of the impact of the rules it helped develop on global public communication. Engaging with external stakeholders, we were told by one team member, “makes me feel validated in the work we’re doing”.⁹⁷ Communicating Facebook’s concerns and the difficulty of developing global rules externally is seen as an important function as well: “They [external stakeholders after engagements] walk away appreciating we’re doing good work”.⁹⁸

At no point in the study did we notice explicit reference to economic interests, much less controlling non-normative interests in the process.

⁹⁰ Observation # 103, 227 documentation of the embedded phase

⁹¹ Observations # 227, documentation of the embedded phase.

⁹² Mark Zuckerberg, “Standing For Voice and Free Expression”, Speech at Georgetown, 17 October 2019, <https://newsroom.fb.com/news/2019/10/mark-zuckerberg-stands-for-voice-and-free-expression>.

⁹³ Ibid. (our emphasis).

⁹⁴ Ibid.

⁹⁵ Observation # 204, documentation of the embedded phase.

⁹⁶ Ibid.

⁹⁷ Observation # 226, documentation of the embedded phase.

⁹⁸ Observation # 228, documentation of the embedded phase.



4.3. Communicative Practices

The process of communication is shaped by the organizational setting that has been implemented and that includes different working group meetings, stakeholder engagement, and input from internal research teams (channelling external research results). The input of internal and external stakeholders shapes the progression of normative change and is used to legitimize certain normative choices over others.

In the working group meetings, we observed multiple references to external experts and stakeholders and direct references to how norms were shaped in reaction to these engagements. Within the team, certain topics – such as hate speech – have specific subject matter experts that have a special responsibility of channelling stakeholder input.⁹⁹ Stakeholder Engagement team members in specific regions also impact normative processes.

In many steps of the process, over working group meetings, the leads for specific policy change processes actively sought out input from the group by asking questions such as “Do folks have concerns?”¹⁰⁰ These open-ended questions afforded responses, beyond the presentation of results from stakeholder engagement for more general inputs based on team members’ understanding of the context of their particular roles – for example, Content Policy team members thinking about the cohesion and the fit between policies, Community Operations team members thinking about enforcement of policies, Public Policy team members thinking about local country conditions and relations with governments. Furthermore, concerns raised were usually not tied directly to either national laws or international norms nor to Facebook’s values such as voice.¹⁰¹

Facebook’s internal research team is responsible for distilling current research and providing context for a proposed rule change. The selection of sources and the presentation is the responsibility of the research team. Their goal, as it has been described to us, is to “distil what we learned so it can shape policy development”. The research team is aware of the difficulty of ensuring a fair representation of non-“Western” sources: “We need expertise that will give us a defensible solution.”¹⁰²

Disregarding one region was considered, in one noted instance, as a noticeable gap: “Was Africa okay with this?”, one working group member asked referring, as we understood it, to both input from stakeholders in Africa and the Stakeholder Engagement team members working in and on Africa.¹⁰³ We observed a personal responsibility, a sense of “geo-thematic ownership”, that team members exhibited regarding “their” stakeholders and/or their topics. The very positive atmosphere observed among the members of the Stakeholder Engagement team will also

⁹⁹ Observation # 480, documentation of the embedded phase.

¹⁰⁰ Observations # 343, 367, documentation of the embedded phase.

¹⁰¹ This raises the question, not addressed here, in how far the selection of stakeholder team members is already a determinative step in the normative outcome.

¹⁰² Observation # 415, documentation of the embedded phase.

¹⁰³ Observation # 341, documentation of the embedded phase.



have an impact on ensuring that, on a purely personal level, all other members will attempt to include the voices of other members (and thus their regions and topics) in the process.

Most engagement with stakeholders happens through private conversations via video or occasionally in face-to-face meetings or workshops: “The heart of our approach to engagement is private conversations, most often in person or by video-conference. We’ve found that this approach lends itself to candid dialogue and relationship-building.”¹⁰⁴ Facebook usually does not disclose the individuals and groups it talks to, arguing that “the conversations can be sensitive, and we want to ensure open lines of communication. Some stakeholders may also request or require confidentiality, particularly if media attention is unwanted or if they are members of a vulnerable community.”¹⁰⁵

¹⁰⁴ Ibid.

¹⁰⁵ Ibid.



5. DISCUSSION: RULE PRODUCTION AS ORDERING

5.1. “Making Rules Up”

This study is based on a combination of normative and empirical research, it applies mixed methods to assess and contextualize various sources of information and benefits from unprecedented access to the product policy process within Facebook, Inc. Though only a pilot case study that provides little bases for generalization, our study has been essential in identifying gaps in current research in terms of understanding the actor constellations, frames of relevance and communicative practices of private norm-development processes that, by virtue of them impacting a global communicative space used by 2.7 billion people, have an important public function.

As one Stakeholder Engagement team member said, with some degree of self-criticism: “We are making rules up”¹⁰⁶. Given the importance of those rules we would like to finally discuss what research question should – based on our results – be addressed to better understand private norm-setting in the context of private communicative service providers.

5.2. Constellations of Actors: Aiming for Throughput Legitimacy

With regard to the constellations of actors of the product policy development process we find that in a non-systematic but representativity-oriented selection process, representatives from different stakeholder groups, particularly civil society and academia, are selected to form part of an increasingly stable actor constellation.

Stakeholder engagement does not appear to us as a defensive practice but, rather, as an epistemic practice motivated by the desire to develop “better” rules. Though we did not observe that economic considerations played a major role in the product policy development processes, their goal – to make Facebook a more attractive platform where the everyone’s “voice” can be heard – is supported by the economically sensible goal to be used more broadly. We also cannot exclude that, even if economic considerations were never voiced during the observed meetings and in our interviews, they have been previously internalized by team members or discussed in fora we did not attend.

The role of leadership – essentially that of Mark Zuckerberg and Sheryl Sandberg – in normative development is a difficult one to clearly establish. While in the internal documents that were shared with us, seeking leadership feedback was one of the steps in the normal progression of a policy development (and some policy changes can entirely originate from leadership), it was difficult to generally establish at which point leadership feedback would be given. We found

¹⁰⁶ Observation # 639, documentation of the embedded phase.



that this depends on the normative change processes, more specifically, their “political” or “technical” nature. At the very least the results of Public Policy Forums are regularly “surfaced up” to leadership to ensure that they have “full visibility”, as one participant put it.¹⁰⁷ The role of integrating leadership feedback here seems to rest, as has been described in a number of journalistic pieces, with Monika Bickert, the VP for Global Policy Management, who would, if needed, “take it [i.e. the issue] to Mark [Zuckerberg]”.¹⁰⁸ We could not shed more light on this. Our methods of observation unfortunately had their limits.

Engaging academics seems to be a very useful approach for the actors involved to increase the epistemic legitimacy of a normative process. The selection of academics is essential to the integrity of this process. Selection can be controlled by consciously engaging with stakeholders representing conceptions other than “Western”, first-generation-rights-focused approaches to human rights, including scholars and activists from African and Arab states, but would also lead to a more challenging stakeholder engagement process as the opinions on policy developments would show much more variety. In any case, even among activists there is a global bias towards “easier-to-realize” rights, such as freedom of expression, while “more-difficult-to-realize” rights, such as the right to health, may be systematically neglected in stakeholder engagements. Algorithmic content curation might only enhance this because first-generation rights are usually more easily “machine-readable” than second- and third-generation rights.¹⁰⁹ As far as academics are concerned, it might be helpful to reflect on their role in this process.

Facebook cannot engage with all of its users and, therefore, selects certain stakeholders who it assumes will be “informed about and able to speak on behalf of others”. Facebook’s selection of stakeholders is, therefore, an exercise of assigning representation. By identifying certain stakeholders as being “able to speak on behalf of others”,¹¹⁰ they make a value-laden choice. In its Stakeholder Engagement policies, Facebook offers justification for its selection of engagement partners by pointing out that the “primary focus of our engagement is civil society organizations, activist groups, and thought leaders, in such areas as digital and civil rights, anti-discrimination, free speech, and human rights.”

While the representativity of civil society organizations (CSO) as agents for the “global good” has been recently seen with more granularity,¹¹¹ and they are sometimes subject to (political) capture, Facebook tries to select specific high-reputation CSOs with a long history of public

¹⁰⁷ Observation # 352, documentation of the embedded phase.

¹⁰⁸ Observation # 651, documentation of the embedded phase.

¹⁰⁹ CJEU, Judgment of 3 October 2019, Case C-18/18 *Eva Glawischnig-Piesczek v Facebook Ireland Limited* at 12.

¹¹⁰ *Ibid.*

¹¹¹ Klonick (2018), 1655; Balkin, Jack M., “Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation”, *University of California, Davis Vol. 51* 2018 (1149-1210) at 1190; Suzor, Nicolas, “Lawless, the secret rules that govern our digital lives”, 2019, at 9 and 92 et seq. arguing that “real change will require the active participation of a broad range of civil society groups, activists, journalists, academics, and regulators. It will be hard work and require many difficult public debates with no easy answers, but there is a great deal at stake.” Black, Julia, “Constructing and contesting legitimacy and accountability in polycentric regulatory regimes”, *Regulation & Governance* (2008) 2, 137–164 (153 and 156).



engagement. As one employee commented, “we have to make a selection”¹¹². The selection is influenced by past experiences – “we’ve gotten good inputs from them before”¹¹³ – but also by the importance of an individual or group within a certain field, the “thought leaders”.¹¹⁴ Sometimes, these are then asked to suggest potential additional stakeholders – “we ask them: to whom should we talk?”¹¹⁵ By engaging with many different stakeholders, by having dedicated regional experts, and by having thematic experts that deal with, for example, hate speech, with violence, or terrorism, Facebook aims at building internal expertise with the aim of controlling selection bias.

As stakeholders that were helpful in the past are consulted again, the stability of the actor constellation increases over time. The definition of the involved actors in the Stakeholder Engagement principles (the “civil society organizations, activist groups, and thought leaders, in such areas as digital and civil rights, anti-discrimination, free speech, and human rights”) influences Facebook’s practice. Digital rights are just human rights applied to technologically mediated settings, civil rights are part of human rights (even if “civil rights” have a special implication in US legal discourse in light of the “Civil Rights Era”) and non-discrimination (a key aspect of “civil rights” as understood in US legal discourse) and free speech are also human rights. This would seem to focus the statement on a commitment to groups and individuals with an expertise in human rights.

While all conversations we observed were held with engaged and interested stakeholders that demonstrated creativity and a richness of ideas in their contributions to the discussion, it was not immediately clear to us how Facebook controls this bias towards human rights-oriented expertise in stakeholder engagement. Yet, as biases go, a bias towards human rights is definitely not *prima facie* problematic. It may become so if the stakeholders whose engagement is sought seem to be primarily influenced by certain human rights traditions or conceptions, primarily those preferring individual rights-conceptions over community-oriented ones.¹¹⁶ This would tend to make policy change processes more likely to lean towards a tendency to over-emphasize “liberal” human rights and underemphasize social and solidarity rights.¹¹⁷

5.3. Frame of Relevance: Autonomously Constructing a Legal Order

Facebook has been constructing a prima facie autonomous and private normative order for public communication that seeks to reconcile interests within that order and is conceived largely without reference to state law or international human rights standards.

¹¹² Observation # 103, documentation of the embedded phase.

¹¹³ Observation # 103, documentation of the embedded phase.

¹¹⁴ Observation # 103, documentation of the embedded phase.

¹¹⁵ Observation # 103, documentation of the embedded phase.

¹¹⁶ Observations # 324, 336 documentation of the embedded phase.

¹¹⁷ All human rights are universal, indivisible and interdependent and interrelated. Cf. Vienna Declaration and Program of Action, adopted by the World Conference on Human Rights in Vienna on 25 June 1993 at I no. 5, <https://www.ohchr.org/en/professionalinterest/pages/vienna.aspx>.



Facebook Public Policy team members confidently participate in creating a separate and unique private normative order for public communication and regulate “novel” questions. They do not directly link this order to any national or international legal order or refer to international human rights commitments. This is, as such, already noteworthy since it is the *Product Policy* team: the normative order has become part of, or even their principal, end product – a normatively bordered socio-communicative space which puts national legal orders before a completely new challenge. This conclusion alone is important because it has implications on the approach taken by existing legal systems in the integration of this private order.

Normative connectors between the internal order and external legal orders – shared norms – exist, but they are rare. While inclusion of civil rights in the US legal tradition, especially equality-related rights and the First Amendment with its very robust protection of free speech, or “voice”, have a solid history in Facebook’s normative processes (including public Civil Rights Audits), the mainstreaming of international human rights is still in its infancy. This is notable, most prominently, in the professed conviction of Product Policy team members that they were “doing good work”.¹¹⁸ They sincerely believe they are doing good work, providing thoughtful standards and trying to engage constructively with multiple stakeholders who themselves are made to appreciate, as one employee put it, they were contributing to “good work” themselves. The “work” they are doing is, our case study found, the construction of an increasingly intricate normative order that will in time be given at least an element of publicness by the introduction of the Oversight Board. The public-private-co-design of this normative order will be the topic of our next case study.

Product Policy team members, in our assessment, tend to understand their role as active participants in the process of developing and refining the shared frame of relevance and explain and justify normative choices or preferences with references to existing policies. Simplicity – to avoid cognitive overload for front-line content moderators – and consistency seemed to us to be important and unwritten normative frames. They also appeared to influence the criteria contained within the November 2019 *Policy Prioritization Framework*.¹¹⁹

That we did not notice any reference to economic interests, much less controlling non-normative interests in the process, does of course not mean that they are irrelevant. It could be suggested that these economic interests (such as a preference for a norm that is best for the bottom line) are ingrained in each single one of the more than fifty individuals closely working on each normative change; but given the data we collected in the case study and the interviews we organized with team members this does not seem to be the case.

Of course, it may be argued that all decisions are influenced by the long-term goal of ensuring the continued attractiveness of Facebook as a social space, which would be the premise of continued commercial success. In that reading any attempt to make the rules more human rights-sensitive (if such sensitivity can be considered a general normative stand-in for the space’s attractiveness) would be influenced by commercial interests. While we do not doubt

¹¹⁸ Observation # 204, documentation of the embedded phase.

¹¹⁹ See section 3.1., above.



that the sustainable success of the “business” represents a significant yet unstated force in all product policy developments, we did not observe a single case where commercial aspects were used as an argument for or against a specific change.¹²⁰

5.4. Communicative Practices: Planned Engagement

The communicative practices are highly proceduralized. The product policy development process, including the stakeholder engagement phase, is a multi-step process with clear timelines, notice periods, possibilities for comments, engagement time, policy formulation, and adoption phases.

In any normative-social setting it holds true that, if the outcome of a procedure might be – for any reason – not intrinsically legitimate, then the proceduralization can increase its legitimacy and make a norm(ative change) amenable to those not agreeing with the particular policy outcome as well. Facebook recognizes this in its Stakeholder Engagement principles: conflicting opinions of stakeholders are not “necessarily” to be reconciled and the whole “spectrum of opinion and points of disagreement” is seen as an enriching source of stakeholder input. Finally, “[n]ot everyone will agree on where we draw the lines, but at a minimum, we need to understand the concerns of those who are affected by our policies”.¹²¹

This is proceduralized internally as well: “We owe internal stakeholders’ responsiveness on this”, one participant said in a meeting,¹²² thus enforcing an understanding of diverse viewpoints, also internally, to be at least considered in normative change processes. However, we also observed that frames of reference were internalized to such a degree as to make it difficult to gauge their actual impact. We found that the key lens through which to analyze the normative process within Facebook was a focus on the communicative practices within the organization.

Facebook employees, even those who had – because of their personal background – undeniable experience with human rights documents, including their enforcement, were not observed to refer to concrete human rights norms during discussions in working groups or during stakeholder engagements.

The research team is very aware of the “Western reading”¹²³ of their standards, we were told, and take particular care to give non-Western regions and their stakeholders “visibility”.

¹²⁰ We note here, again, that we focused on the Product Policy development and not the policy development process related to advertising. As the changes to Ad Policy are also decided in the Public Policy Forum, the individuals within the Stakeholder Engagement team heading the processes overlap and the processes of policy development sometimes run parallel to those related to changes in the Product Policies, we cannot exclude a certain influence which we, however, did not study in depth.

¹²¹ Facebook, Stakeholder Engagement, https://www.facebook.com/communitystandards/stakeholder_engagement.

¹²² Observation # 254, documentation of the embedded phase.

¹²³ Observation # 427, documentation of the embedded phase.



Future research should aim at a better understanding of how the process structures influence norm production and, more specifically, what new practices will develop to integrate the Oversight Board as a new key actor.

5.5. Outlook

The run-time of the project has fallen into a phase when Facebook introduced a supervisory body made up of independent experts that can take (mostly) binding decisions in certain content-related disputes. In more ways than one, developing the Oversight Board will impact global approaches to private rulemaking for public communication. While Facebook seems to consider case-based decisions as fundamentally different to planned policymaking we believe that this is not a given but that it will rather be an interesting empirical question as to how board decisions influence policy development processes.

Wishing to gather “global feedback” on the construction of the Oversight Board, Facebook began a global consultation process, which allowed users (technically everyone) to submit views directly to Facebook. As the Stakeholder Engagement team writes in their principles: “One could imagine a similar process whereby NGOs and civil society organizations could join our network of contacts in order to receive regular policy updates and provide feedback to members of our team.”¹²⁴ Might this be the start of broader consultation on (even) the meta norms that structure the policy development process?

The results presented in this paper show how one dominant social media company frames the make-up of its communicative space. It did so (and still does) by constructing an order that has, apart from some anchors in US law, no firm foundation in any one national legal order. Product policies are intricately interlinked with the goals of the communicative space Facebook provides. As a semantic hint, it is the *Product* Policy team that develops new policies. In this reading, it can be argued that Facebook’s *product* is, therefore, the socio-communicative space it provides to the public including the communicative infrastructure and the normative form the product policies take in the form of their Community Standards.¹²⁵

This can be seen as necessary normative answer to the privatization of spaces for public communication (where communication goes, rules for communication must follow). But regardless of the perspective we have to better understand how the normative process works, given the impact this private ordering approach has on individuals’ spheres of communicative freedom and the social cohesion of our societies. This case study represents an initial step into that direction. Analyzing the private-public co-design and administration of the Oversight Board and its impact on Facebook’s private order of communication and the public order will be our

¹²⁴ Ibid.

¹²⁵ We are aware, though, of other readings of what Facebook’s “product” is. “We run ads”, Mark Zuckerberg memorably reminded a US Senator during a Congressional hearing. Ads and advertising revenue are based on the attention they are given by Facebook users. Arguably, therefore, Facebook’s “product” is its users’ attention, which is provided to ad buyers.





next task. If the rules are the product, this might mean that the *product* will in future be co-designed by wider society – assuming that the members of the Board frame their role as “doing good” in the sense of serving the “common good” and not just interpreting the private orders set by Facebook.